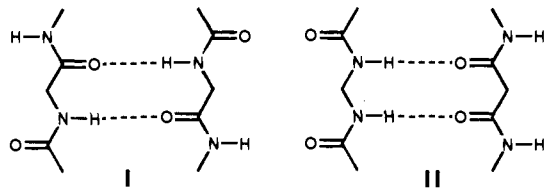
Evaluation of the Conformation-Directing Effects of Secondary Hydrogen-Bonding Interactions in Flexible Tetrapeptide Analogues

Robb R. Gardner and Samuel H. Gellman*

Department of Chemistry, University of Wisconsin Madison, Wisconsin 53706

Received June 12, 1995

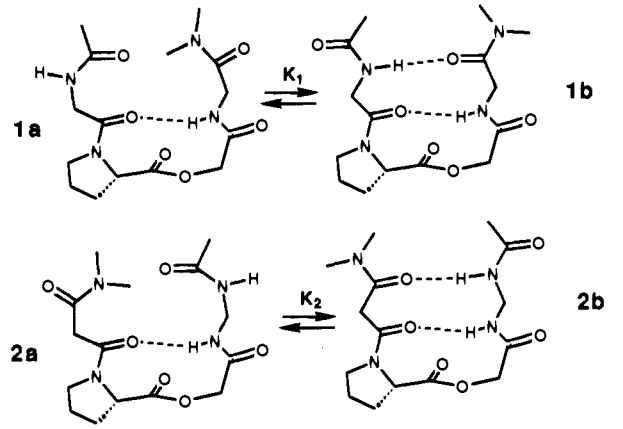
Hydrogen bonds are involved in many inter- and intramolecular recognition processes, but simply counting the hydrogen bonds in a complex or folding pattern does not necessarily provide insight on stability.[1] In 1990, Jorgensen and Pranata pointed out that variations among the association constants for a series of triply hydrogen bonded complexes, including the guanine−cytosine base pair, could be rationalized in terms of "secondary interactions" among the hydrogen-bonding groups.[2] These secondary interactions involve electrostatic attractions or repulsions between donor (electropositive) and acceptor (electronegative) atoms forced to approach one another in the course of forming the primary hydrogen bonds. The predictive value of the secondary interaction hypothesis has been demonstrated experimentally with conformationally rigid heterocycles.[3−5] In their original paper, Jorgensen and Pranata extended this concept to peptides and peptidomimetics: the hydrogen-bonded dimer of a diamide derivative of glycine (**I**) was calculated to be substantially weaker than the dimer of diamide derivatives of diaminomethane and malonic acid (**II**), when the hydrogen-bonding partners were constrained to be planar.[2] Described here is an experimental comparison of hydrogen-bonding patterns **I** and **II**.[6]



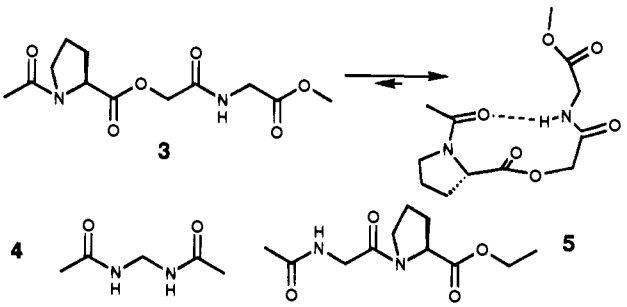
I **II**

Our approach involved examination of intramolecular hydrogen bond formation in isomers **1** and **2** and reference compounds **3**−**5**, via N−H stretch region IR data obtained in dilute CH_2Cl_2 solution.[7] Precedent suggested that the folding behavior of **1** and **2** in nonpolar solvents would conform largely to the two-state equilibria shown in Scheme 1.[8,9] For **3**, which

Scheme 1



represents the central prolyl-glycolyl segment common to **1** and **2**, the 10-membered-ring hydrogen-bonded folding pattern is predominant, as indicated by the major band at 3318 cm^{-1} (Figure 1). The minor band at 3430 cm^{-1} arises from a small population of N−H free of intramolecular hydrogen bonding. Diamide **4** displayed only one N−H stretch band, at 3446 cm^{-1} (not shown), indicating that there is no six-membered-ring hydrogen bonding within the diacyl diaminomethane residue.



3 **4** **5**

For both **1** and **2**, there is a major band for N−H hydrogen bonded to an amide carbonyl (ca. 3330 cm^{-1}, Figure 1).[10] Control studies indicated that this hydrogen bonding is strictly intramolecular.[11] In addition, **2** displays a significant band at 3441 cm^{-1}, and **1** displays a shoulder at 3404 cm^{-1}. The additional band for **2** was assigned to the C-terminal N−H group in a non-hydrogen-bonded state, on the basis of examination of the version of **2** in which the C-terminal amide group was labeled with ^{15}N (ca. 7 cm^{-1} shift to lower energy for the non-hydrogen-bonded N−H stretch band).[11,12] The additional band for **1** was assigned to the N-terminal N−H on the basis of examination of **5**, which displays a major band at 3410 cm^{-1} and a minor shoulder at 3448 cm^{-1}. Precedent indicates that the major band at 3410 cm^{-1} for **5** results from the so-called C_5 interaction, a weak intraresidue C=O···H−N attraction;[13]

(1) For examples in which folding patterns containing one intramolecular amide−amide hydrogen bond are more enthalpically stable than folding patterns containing two intramolecular hydrogen bonds, see: (a) Liang, G.-B.; Rito, C. J.; Gellman, S. H. *J. Am. Chem. Soc.* **1992**, *114*, 4440. (b) Dado, G. P.; Gellman, S. H. *J. Am. Chem. Soc.* **1993**, *115*, 4228.

(2) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008. See also: Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810.

(3) Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010. See also: Murray, T. J.; Zimmerman, S. C.; Kolotuchin, S. V. *Tetrahedron* **1995**, *51*, 635.

(4) Jeong, K. S.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Famulok, M.; Rebek, J. *J. Am. Chem. Soc.* **1991**, *113*, 201.

(5) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 209.

(6) For use of malonic acid and diaminomethane derivatives in "retro-inverso" peptidomimetics, see: Goodman, M.; Chorev, M. *Acc. Chem. Res.* **1993**, *26*, 266 and references therein.

(7) For interpretation of N−H stretch region IR data in CH_2Cl_2, and examples of quantitative analysis of N−H stretch data, see: (a) Dado, G. P.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 1054 and references therein. (b) Nowick, J. S.; Abdi, M.; Bellamo, K. A.; Love, J. A.; Martinez, E. J.; Noronha, G.; Smith, E. M.; Ziller, J. W. *J. Am. Chem. Soc.* **1995**, *117*, 89 and references therein. (c) Tsang, K. Y.; Diaz, H.; Graciani, N.; Kelly, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 3988 and references therein. (d) Winningham, M. J.; Sogah, D. Y. *J. Am. Chem. Soc.* **1994**, *116*, 11173.

(8) Boussard, G.; Marraud, M.; Neel, J.; Maigret, B.; Aubry, A. *Biopolymers* **1977**, *16*, 1033.

(9) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4105.

(10) ^1H NMR data for **1** and **2** in $CDCl_3$ indicate that in each case there is a very small amount (<5%) of the *cis*-proline rotamer.

(11) Data may be found in the supporting information.

(12) For a localized A−B stretch, the band position can be estimated from the equation $\nu = (2\pi c)^{-1}[k(M_A + M_B)/M_AM_B]^{1/2}$, where c is the speed of light, k is the force constant of the A−B bond, M_A is the mass of atom A, and M_B is the mass of atom B. (Silverstein, R. M.; Bassler, G. C.; Morrill, T. C. *Spectrometric Identification of Organic Compounds*, 5th ed.; John Wiley & Sons: New York, 1991; p 93.) This calculation predicts a localized ^{15}N−H stretch to be ca. 12 cm^{-1} lower in energy than a ^{14}N−H stretch.

(13) (a) Avignon, M.; Huong, P. V.; Lascombe, J.; Marraud, M.; Neel, J. *Biopolymers* **1969**, *8*, 69. (b) Burgess, A. W.; Scheraga, H. A. *Biopolymers* **1973**, *12*, 2177.

0002-7863/95/1517-10411$09.00/0

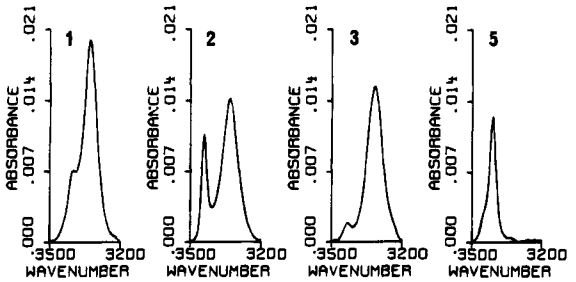


Figure 1. N−H stretch FT-IR data for 1 mM samples in CH_2Cl_2 at room temperature, after subtraction of the spectrum of pure CH_2Cl_2 (nominal resolution = 2 cm^{-1}). From left to right: **1**, maxima at 3404 (shoulder) and 3328 cm^{-1}; **2**, maxima at 3441 and 3332 cm^{-1}; **3**, maxima at 3430 and 3318 cm^{-1}; **5**, maxima at 3448 (shoulder) and 3410 cm^{-1}.

the minor shoulder arises from non-hydrogen-bonded N−H. These data suggest that the shoulder at 3404 cm^{-1} for **1** arises from the N-terminal N−H involved in a C_5 interaction. This conclusion is supported by the observation that **1** with N-terminal [^{15}N]Gly has the shoulder shifted to 3399 cm^{-1}, while **1** with C-terminal [^{15}N]Gly has the shoulder at 3406 cm^{-1}.[11,12]

The IR data discussed above indicate that the two-state conformational equilibria shown in Scheme 1 are good descriptions of the behavior of **1** and **2** in dilute CH_2Cl_2 solution. In order to estimate equilibrium constant K_1, we used the data for **5** and curve-fitting analysis to estimate the amount of N−H contributing to the 3404 cm^{-1} shoulder in the spectrum of **1**.[9,11] We assume that the concentration of this type of N−H group corresponds to the concentration of folding pattern **1a**, and that the remainder of the total concentration of **1** corresponds to folding pattern **1b**. According to this analysis, $K_1 = 0.67 \pm 0.13$. A similar analysis, with use of **4** to generate the extinction coefficient characteristic of the non-hydrogen-bonded C-terminal N−H in folding pattern **2a**, gives $K_2 = 0.44 \pm 0.05$. The accuracy of these values is limited by the assumptions underlying each analysis (indeed, visual inspection of the IR data for **1** and **2** suggests a larger difference between K_1 and K_2 than is manifested in the calculated values). With regard to the role of secondary interactions in the folding of **1** and **2**, it seems most conservative to conclude that K_1 and K_2 are indistinguishable.

The similar propensities of **1** and **2** to adopt doubly hydrogen bonded conformations (**1b** and **2b**) in methylene chloride contrast with the computational findings for planar intermolecular complexes **I** and **II**.[2] The stability difference calculated for planar **I** vs planar **II** (each relative to isolated components) was nearly identical to that for guanine/cytosine vs uracil/2,6-diaminopyridine; experimental association constants for complexes related to these latter two differ by ca. 100-fold in chloroform.[2] The difference between the calculated result for **I** vs **II** and the data for **1** vs **2** probably arises, at least in part, from the computational assumption of planarity for the constituents of **I** and **II**. In the real molecules, intramolecular dipole−dipole repulsions are likely to promote nonplanar conformations within each diamide unit, particularly for derivatives of diaminomethane and malonic acid (the constituents of **II**).[14]

We have provided the first experimental evaluation of secondary hydrogen-bonding interactions in flexible systems. Our data show that the optimal arrangement of secondary interactions between two conformationally mobile arrays of donor and/or acceptor groups does not necessarily correspond to the most stable hydrogen-bonding pattern. In contrast, analysis of secondary interactions constitutes a powerful tool for evaluating multiply hydrogen bonded interactions in conformationally rigid systems.[2-5] A key distinction between rigid and flexible arrays of the hydrogen-bonding groups is that dipolar repulsions between rigidly linked donors or acceptors are inescapable, but such repulsions can be avoided in flexible structures via torsional mobility.[15]

Supporting Information Available: Variable concentration ^{1}H NMR data for **1** and **2** (aggregation control studies) and IR data pertaining to the determination of K_1 and K_2 (9 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA951897G

(14) (a) Stern, P. S.; Chorev, M.; Goodman, M.; Hagler, A. T. *Biopolymers* **1983**, *22*, 1901. (b) Tereshko, V.; Navarro, E.; Puiggali, J.; Subirana, J. A. *Macromolecules* **1993**, *26*, 7024.

(15) This research was supported by the National Science Foundation (CHE-9224561). S.H.G. thanks the NSF-PYI program (CHE-9157510), Procter & Gamble, and Merck Research Laboratories for support. We thank Prof. James Nowick for helpful comments.